(4)

VLSI Memo No. 89-513
March 1989

Equipment Models for Process Optimization and Control Using Smart Response Surfaces

E. Sachs, G. Prueger, and R. Guerrieri

Abstract

An equipment model has been developed for the low pressure chemical vapor deposition (~~LPCVD~~) of polycrystalline silicon in a horizontal tube furnace. The model predicts the wafer-to-wafer deposition rate down the length of the tube. Inputs to the model include: silane flow rates from three injectors, injector locations, locations of and temperatures of three thermocouples, operating pressure, the number of wafers, wafer diameter, the location of the wafer load, and other physical dimensions of the furnace such as tube length, and inner diameter. The model is intended to aid the process engineer in the operation of equipment, and the equipment designer in the design of new equipment.

The onc dimensional finite difference model encompasses the convective and diffusive fluxes of silane and hydrogen in the annular space between the wafer load and tube walls. The reaction of silane is modeled, including the generation and transport of hydrogen. Kinetic and injection parameters in the model were calibrated using a series of nine statistically designed experiments. The model accurately predicts the axial deposition profile over the full range of experimentation and demonstrates good extrapolation beyond the range of experimental calibrations. The model was used to predict a set of process parameters that would result in the least variation of deposition rate down the tube. The predicted parameters agree well with experimentally determined optimum conditions.

Microsystems Research Center Room 39-321 | Massachusetts Institute of Technology | Cambridge Massachusetts 02139 | Telephone (617) 253-8138

Acknowledgements

This research was supported in part by the Defense Advanced Research Projects Agency under contract number N00014-85-K-0213 and by the Microelectronics and Computer Technology Corporation.

Author Information

Sachs: Department of Mechanical Engineering, MIT, Room 35-229, Cambridge, MA 02139. (617) 253-5381.
Prueger, *current address:* Rocketdyne Division, Rockwell International Corporation, Mail Code AC51, 6633 Canoga, Canoga Park, CA 91303 (818) 710-3359.
Guerrieri, *current address:* University of Bologna, Bologna, Italy.

EQUIPMENT MODELS FOR PROCESS OPTIMIZATION AND CONTROL USING SMART RESPONSE SURFACES

E. Sachs (1) G. Prueger (1) R. Guerrieri (2)

[1]Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

[2]University of Bologna
Bologna, Italy

An equipment model has been developed for the low pressure chemical vapor deposition (LPCVD) of polycrystalline silicon in a horizontal tube furnace. The model predicts the wafer-to-wafer deposition rate down the length of the tube. Inputs to the model include: silane flow rates from three injectors, injector locations, locations of and temperatures of three thermocouples, operating pressure, the number of wafers, wafer diameter, the location of the wafer load, and other physical dimensions of the furnace such as tube length, and inner diameter. The model is intended to aid the process engineer in the operation of equipment, and the equipment designer in the design of new equipment.

The one dimensional finite difference model encompasses the convective and diffusive fluxes of silane and hydrogen in the annular space between the wafer load and tube walls. The reaction of silane is modeled, including the generation and transport of hydrogen. Kinetic and injection parameters in the model were calibrated using a series of nine statistically designed experiments. The model accurately predicts the axial deposition profile over the full range of experimentation and demonstrates good extrapolation beyond the range of experimental calibration. The model was used to predict a set of process parameters that would result in the least variation of deposition rate down the tube. The predicted parameters agree well with experimentally determined optimum conditions.

1 INTRODUCTION

Motivation and Background

The semiconductor industry has at its disposal advanced simulation and design tools to aid in such tasks as circuit design and device design. However, there are relatively few tools available to the equipment designer and the process engineer. The goal of this work is to contribute to the development and use of equipment models which can be used to design and operate fabrication equipment.

The current work concerns the low pressure chemical vapor deposition (LPCVD) of polycrystalline silicon in a horizontal hot walled reactor or tube furnace. LPCVD processes have become a widespread family of processes, having largely replaced atmospheric pressure reactors due to the improved product uniformity and reduced defect density possible at low pressures. The reduced pressure increases the diffusivity in the gas leading to better uniformity, while the lower gas density decreases the occurrence of free space reaction and the associated particle contamination.

Due to the high packaging densities required in batch LPCVD processes, product uniformity, especially across a wafer and wafer to wafer, can be problematic. In LPCVD of polysilicon the uniformity of deposition thickness is a key concern, with the largest variation observed from wafer to wafer in a batch. This work presents the development and use of an equipment model which concerns the uniformity of deposited film thickness in the LPCVD of polysilicon.

Equipment Description

The experimental work described in this paper was performed on a BTU Engineering/Bruce System 7351C horizontal, hot wall furnace, which is typical of the commercial reactors used for LPCVD processes. As shown schematically in figure 1, the process area consists of a quartz process tube surrounded by a three zone heating coil with a quartz liner inside the process tube. The wafer load is inserted and removed on a silicon carbide cantilever attached to the front door of the reactor. The wafer load rests in quartz boats and is situated concentrically in the tube so that the wafers are

perpendicular to the main gas flow. The vacuum is pulled from the source[1] end of the tube. Pressure is electronically controlled in the furnace by a butterfly valve in the vacuum gas line. Gas (pure silane, SiH_4) can be injected through three injectors, a fixed load end injector and two moveable injectors, which are usually situated in the center and source zones. The temperature of the zoned heater is controlled during deposition by three thermocouples which are situated in a quartz sheath inside the process tube and are individually located near the beginning of the wafer load, in the center, and near the end of the wafer load.

Related Work

Many workers have investigated the fundamentals of the pyrolysis of silane to form silicon. Hitchman [1] derived a basic linear model of the LPCVD reaction kinetics. Van Den Brekel [2] and Claassen [3] investigated the LPCVD reaction chemistry for polysilicon and developed an understanding of the main effects of the reaction chemistry as supported by experimental observations. Middleman [4] developed a numerical model for the mass transport in an annular LPCVD reactor. He showed the effect of including diffusion along with the gas convection and that the flow in the annular region was insufficient to create significant circulating flows between the wafers.

A numerical model for the LPCVD of polysilicon was developed by Jensen [5, 6]. This model embodies the most advanced reaction kinetic model derived from the studies of Van Den Brekel [2] and Claassen [3] with a gas flow model dependent on the system geometry. The model predicts the axial deposition profile for polysilicon deposition for a ramped temperature processes in which there is no gas injection other than at the load end of the furnace.

An expert system approach was recently developed at Berkeley [7] for the LPCVD of doped polysilicon. The expert system partitions the deposition process goals into six modules for determination of the resistivity, thickness, uniformity, grain size, film stress, and a support module. Each individually searches a data base of empirical rules for the correct equipment settings to meet the process specifications. The expert system is limited to operation in a small operating window in which it has empirical knowledge.

---

[1] Throughout this paper the furnace will be referred to in three zones; the load zone which is the front area of the furnace where the wafers are inserted and removed, the source zone which is the area at the rear of the tube, and the center zone.

2 NEED FOR EQUIPMENT MODELS

Definition and Configuration

In the most general sense, an equipment model is a body of knowledge which provides predictions about the outputs from a unit manufacturing process, given information about the inputs to the process. Figure 2 illustrates a generic equipment model with outputs and two classes of inputs, the process parameters and the disturbances. The process parameters are those parameters that we exercise direct control over, for example, temperature, pressure, and gas flow rates. The disturbances are those inputs to the process which are subject to unintended and undesired variations. In some cases, the magnitude of the disturbances can be monitored, while in other cases they cannot. Examples of disturbances include variations in the properties of incoming materials and variations in the process parameters themselves.

A competent equipment model must provide information about the outputs and the variation of the outputs as a function of the process parameters and the disturbances. The variations might include across the wafer variation, wafer to wafer variation, and batch to batch variation. A model might address all three classes of variation, or might focus on the most important class as indicated by experience. Accurate predictions about the process mean are often less important than accurate predictions about variation, since in most processes the mean can be adjusted to its target value without a substantial effect on variation. An example of such an adjustment would be the length of time in a LPCVD deposition.

Uses for Equipment Models

Equipment models can be used to aid in the operation of existing equipment, or in the design of new equipment. In the area of operations, equipment models can be used to optimize the operation of a process. Typically, the model would be used to find a set of process parameters which result in the least variation of the outputs . If the equipment model includes predictions of the effective disturbances, the optimization procedure using the model can include minimization of the effect of the disturbances. The use of equipment models for optimization is discussed later in this paper.

Equipment models can also be used for process control. Having selected the operating point in the optimization procedure, the model can now serve to guide adjustments made locally around the operating point. For example, as an LPCVD tube drifts with build up from successive runs, the model can be used to predict the changes in gas flow rates needed to bring the results back as close as possible to the target.

A model which captures the effect of internal parameters such as geometric dimensions, and choice of materials, can be used as a simulation tool for the design of new equipment. Such a model can substantially reduce the development time for new equipment.

Construction of Equipment Models

Equipment models may be constructed by two distinctly different approaches: physically based mechanistic modeling and statistical modeling. Each approach has its distinct advantages. Physically based models have the advantage of broad applicability, good extrapolation beyond the range of experimental verification, and good prediction of process sensitivities. Statistical modeling has the advantages of ease of application and good absolute accuracy within the range of measurement.

Physically based models may be either closed form or numerical (finite element methods, boundary element methods) in nature. Statistical models are most effectively developed using techniques of statistical experimental design, such as "factorial experimental design and response surfaces" [9] and Taguchi "orthogonal array" [8]. The unifying feature of designed experiments is that all the parameters of interest are varied simultaneously, in contrast to the more conventional one variable at the time experimental techniques. In this way, the total experimental range is explored with a minimum number of experiments.

Goal of the Current Work

In current practice, the two approaches to model construction discussed above have been followed independently. The broad goal of this work is to fuse the two methods and gain the benefits of both. The model resulting from the combination of experimental design and physical understanding is called a "smart surface".

The current work can be understood by reference to the flow diagram of Figure 3. After accumulating running experience and understanding the process physics, two directions are taken. First, a mechanistically based model is developed. In this work, the model is a finite difference model which has four adjustable coefficients embedded in it which represent areas of uncertainty about the physics. On the parallel path, a series of designed experiment is pertained. The designed experiments are used to optimize the process and to calibrate the mechanistic model, thereby creating a "smart response surface". The smart response surface is then used to optimize the process, and this optimum is compared with the optimum from the experiments themselves.

3 MODEL DEVELOPMENT

Modeling Approach

The model consists of a one-dimensional representation of the LPCVD reactor for polysilicon. The inputs to the model include: the physical dimensions of the process area (i.e. linear diameter, process tube length, injector diameter), number of wafers in the load, diameter of the process wafers, position of the wafer load, flow rates from the three injectors, positions of two of the injectors, and temperatures and positions for each of the profile thermocouples. The axial temperature profile is determined by a linear interpolation of the temperatures for each of the thermocouple sites. The axial flow is modeled with a convection-diffusion representation incorporating area changes due to wafer load, and a laminar, plug flow velocity distribution.

The model predicts the axial deposition profile of the polysilicon. This solution is calculated using a Newton-Raphson method on the center-difference numerical representation of the one-dimensional system.

The following sections describe the analysis behind the assumptions in the model and the manner in which it was constructed.

Process Physics

The process physics can be understood by order of magnitude analysis of the following physical mechanisms:

- introduction of gas into the furnace
- mass transport in the axial flow direction
- mass transport between the wafers
- heat transfer
- chemical reaction at the hot surfaces and depletion of silane

Order of magnitude analysis of the injector gas [5] exit velocities have shown that the velocities are on the order of 0.5 x Mach 1 and that there is a large pressure drop at the exit of the injector. The result of this analysis indicates that there is an expansion wave of the exit jet into the furnace [10]. The physical understanding of this system is limited.

The relative importance of the convective flux to the diffusive flux in the annular area is captured by the Peclet number, $Pe_a \equiv \frac{Vd}{D}$. Where V is the estimated average gas velocity in the annular region ($V \sim 200$ cm/sec), d is a relative length on the order of

the furnace length (d ~ 200 cm), and D is the diffusion coefficient at the operating temperature and pressure (D ~ 6000 cm²/sec). Since this analysis shows the Peclet number to be on the order of one, both the convective and diffusive fluxes must be represented in the model.

An indication of the flow regime can be determined by the Reynolds number, $Re \equiv \frac{\rho V d}{\mu}$ Where V is the estimated average gas velocity in the annular region (V ~ 200 cm/sec), ρ is the gas density ($\rho \sim 2 \times 10^{-4}$ kg/m³), d is a relative length on the order of the furnace diameter (d ~ 20 cm), and μ is the gas viscosity ($\mu \sim 3 \times 10^{-5}$ kg/m sec). The Reynolds number was found to be on the order of one. This indicates that a laminar flow approximation can be made.

A Poiseulle flow analysis of the furnace pressure distribution accounting for the possibility of slip flow at the boundaries [11] indicates that there was a negligible pressure drop down the length of the furnace. This pressure drop was on the order of 0.1% of the total pressure. This meant that a constant pressure could be assumed in the furnace.

The Peclet number can be used to indicate the relative importance of the convective and diffusive terms in the radial direction by changing the relative order of the length to that of the liner diameter (~ 2 cm) and estimating the radial velocity between the wafers due to the reaction stoichiometery producing a net flux of hydrogen out from between the wafers. This analysis showed that the Peclet number was much less than one, ($Pe_r \ll 1$), indicating that convective transport of hydrogen out from between the wafers does not influence the diffusive transport of silane in to the wafer surfaces.

Another method of analysis of the radial mass transport effect can be performed utilizing the Sherwood number, $Sh \equiv \frac{kd}{D}$. Where k is the mass transport coefficient of the reaction from the Arrhenius reaction dependence (k ~ 8 x 10⁻⁸ cm/sec), d is a relative length ON the order of the tube diameter (d ~ 20 cm), and D is the diffusion coefficient (D ~ 6000 cm²/sec). The Sherwood number is also much less than one (Sh « 1). This indicates that growth rate is reaction rate limited, or that the diffusion time for the gases between the wafers is much smaller than the reaction time at the wafer surface, indicating that the radial gas concentration can be considered uniform.

An order of magnitude analysis for the relative importance of radiative to convective heat transfer can be captured for small temperature variations as: $\frac{\sigma T^3 d}{k}$, where σ is the Stefan-Boltzmann's constant (5.67 x 10⁻⁸ W/m² °K⁴), T is the operating temperature (T ~ 898°K) d is a relative length on the order of the tube diameter (d ~ .214m), and k is the thermal conductivity of the gas (k ~ 2 x 10⁻² W/m K). The analysis

show that $\frac{\alpha T^3 L}{k} \sim 42$ indicating that radiation is the dominant mode of heat transfer. This allows for an isothermal assumption in the radial direction.

The reaction kinetics involved in the pyrolysis of silicon from pure silane has been studied [2, 3] and the general consensus has been that the silane, SiH_4 breaks down in the gas phase as:

$$SiH_4(g) \rightleftharpoons SiH_2(g) + H_2(g) \tag{1}$$

The SiH_2 then adsorbs on the hot surfaces, silicon is deposited, and the by-product, hydrogen, desorbs:

$$SiH_2(g) \rightleftharpoons SiH_2(a) \tag{2}$$

$$SiH_2(a) \rightleftharpoons Si(s) + H_2(a) \tag{3}$$

$$H_2(a) \rightleftharpoons H_2(g) \tag{4}$$

The reaction equation for the deposition of silicon from silane described by the above chemical reaction equations and substantiated by experimental evidence was formulated by Roenigk and Jensen [6] as:

$$\mathcal{R} = \frac{k_1 X_{SiH_4} C_{tot}}{1 + K_s X_{SiH_4} C_{tot} + K_h \sqrt{C_{tot}\,(1 - X_{SiH_4})}} \left[\frac{\text{molSi}}{\text{m}^2\text{sec}}\right] \tag{5}$$

Where X_{SiH_4} is the molar fraction of silane, $(1 - X_{SiH_4})$ represents the molar fraction of hydrogen, C_{tot} is the total concentration of the gases, K_s and K_h are the equilibrium coefficients (adsorption to desorption) of silane and hydrogen and indicate the effect of silane and hydrogen concentrations on the silicon deposition, and k_1 is an Arrhenius reaction rate dependence represented as:

$$k_1 = k_0 \exp\left(\frac{\Delta E_A}{RT}\right) \left[\frac{\text{mol}}{\text{m}^2\text{secatm}}\right] \tag{6}$$

Where k_0 is the Arrhenius reaction constant, ΔE_A is the activation energy, R is the universal gas constant, and T is the temperature. This determines the effect of temperature on the reaction.

The depletion of silane and generation of the hydrogen by-product is the cause of the wafer-to-wafer non-uniformity in the furnace. The depletion of silane for 150 std. cm^3/min of silane injection and a system of 150 six inch wafers with an average growth rate of 50 Å/min is on the order of 40%.

Model Formulation

The LPCVD of polysilicon is a reaction rate limited process, Sh « 1. Due to the low pressures, diffusion coefficients are high giving the reactant species ample time to diffuse radially to the wafer surfaces. The problem in a horizontal tube furnace is that the geometry creates an environment in which wafers at different axial positions, although at equal pressure, see different concentrations of the reactant species, silane, due to depletion and by-product, hydrogen, generation. Close wafer to wafer uniformity is difficult to maintain. A one-dimensional model is appropriate to analyze this system.

The annular Peclet number ($Pe_r \sim 1$) indicates that the model must account for the convective and diffuse fluxes of the gas species in the axial, or z, direction. Since the reaction model only depends on the concentration of silane and hydrogen, the system can be considered binary.

The gas injection is incorporated in the model through an empirical function and is included in the silane flux balance as a silane generation term. The amount of silane generation at each axial location in the furnace is determined by an empirical injection function. The length of injection is determined by a linear relationship with the injector flow:

$$l = con\Delta Q_{injected} \tag{7}$$

Where $Q_{injected}$ is the flow rate of the gas from the injector, con is an adjustable parameter, Δ is the wafer spacing, and l represents the effective distance of the gas spray.

Two empirical functions were proposed to model the injection flow: a flat model, Figure 4a, and a ramp mode, Figure 4b. The ramp model is more physically representative because it incorporates the higher concentration of gas at the injector exit, which would result from an expansion wave. The flat injector model gives an average approximation of the exit jet phenomena. Choice of the best injector spray model will be discussed in Section 5.2.

Figure 5 gives a schematic representation of the one-dimensional LPCVD polysilicon model. The wafer load is represented by a block mass with a specified reaction area per unit length, L_r, as determined by the total wafer area, the wafer spacing, and the tube surface area.

The model equations consist of the convection diffusion flux equation for silane:

$$-\frac{d}{dz}\left(A_{tw}DC_{tot}\frac{dX_{SiH_4}}{dz}+A_{tw}C_{tot}\frac{d\psi}{dz}X_{SiH_4}\right)=-(R-F(z)) \tag{8}$$

Where the first term on the left hand side is the diffusion term and the second term accounts for convection. On the right hand side, R accounts for silane depletion due to reaction at the wafers and tube wall and $F(z)$ is a silane generation term defined by the injector function.

Coupled with this equation is the sum of the hydrogen and silane fluxes giving the mass conservation equation:

$$-\frac{d}{dz}\left(A_{tw}C_{tot}\frac{d\psi}{dz}\right)=(R+F(z)) \tag{9}$$

Where A_{tw} is the cross sectional flow area defined according to the position in the tube and

$$-\frac{d\psi}{dz}=\mathbf{V} \tag{10}$$

is a potential flow representation of the molar average velocity, V, used for numerical efficiency. The diffusion terms have cancelled since this is a binary system, leaving the convective term for the gas. The stoichiometry of the chemical reaction, in which for every molecule of silicon deposited two molecules of hydrogen are generated, has been accounted for as can be seen by the sign of the reaction term, R. $F(z)$ is as stated above.

The boundary equations associated with Equations 8 and 9 are:

$$-\left(A_{tw}C_{tot}D\frac{dX_{SiH_4}}{dz}+A_{tw}C_{tot}\frac{d\psi}{dz}X_{SiH_4}\right)_{z=0}=Q_0 \tag{11}$$

$$\left(\frac{dX_{SiH_4}}{dz}\right)_{z=L}=0 \tag{12}$$

$$-A_{tw}\,|_{z=0}\,C_{tot}(z=0)\left(\frac{d\psi}{dz}\right)_{z=0}=Q_0 \tag{13}$$

and

$$\psi(L)=0 \tag{14}$$

Where Equations 11 and 13 are consistent with the assumption that the gas flow and total velocity at the load end of the furnace is determined by the amount of gas injected

through the load injector. Equation 12 states that the concentration is constant at the source end of the furnace. This can be assumed since there is very little reaction at the source end of the furnace. Equation 14 gives the reference for the potential representation of the gas velocity.

Numerical Technique

The discretization of the equations consists of a center difference method over a one dimensional grid which consisted of N+1 grid points having coordinates z_i, $i = 1$, N+1. Equation 8 and 9 were decoupled and then linearized using the Newton-Raphson method. The resulting discretized equations were arranged as tri-diagonal matrices and solved using a standard solving package.

Adjustable Parameters

There are four adjustable parameters in the model which will be fitted to experimental data collected from a designed experiment.

Three of the constants are in the reaction model of Equation 5. They are k_1, K_s, and K_h. These parameters are physical constants that arise from the chemical reaction for the deposition of polysilicon. The Arrhenius dependency of the reaction rate is found in k_1, where k_0 from Equation 6 is to be extracted from experimental data. The effects of the silane and hydrogen adsorption and desorption on the growth rate are quantified by K_s and K_h, respectively.

The fourth constant is in the empirical injector function for the spray length, Equation 7.

4. EXPERIMENTATION

Experiment Design

A set of experiments was designed to calibrate and test the accuracy of the model. The Taguchi L_9 orthogonal array, developed by Genechi Taguchi [8], was chosen for this experimental design, Figure 6. This design allowed the investigation of a large operating space, with four equipment parameters at three levels, in only nine experiments. The limitations of this design are that the four parameters would have to be independent of each other and that factor interactions could not be studied from the results. These are not inhibiting limitations. Independent factors can always be found or interacting factors can be combined into one independent factor [12].

Before the equipment parameters to be used as experiment factors where chosen it was necessary to determine the effect of equipment disturbances on the deposition profile. A group of parameter settings was found which gave a relatively flat profile. These settings were the base line equipment settings. The furnace was run with these settings for a number of runs which spanned approximately 100,000 Å of deposition on the furnace walls. After a clean of the equipment, the base line was run again. This set of experiments indicated that the repeatability of the equipment was not greatly effected by equipment disturbances. Figures of the measured data include error bands for the experimental growth rates as determined from these experiments.

The choice of equipment parameters to be used as experimental factors was important. The model was developed to predict the axial deposition profile. To verify the model, it was necessary to choose equipment parameters which had a dominant effect on this profile. The factors chosen were:

- pressure
- load injector flow rate (Q_{load})
- center injector flow rate (Q_{center})
- position of source injector (X_{source})

Factor levels for the experiment factors were chosen based on the base line settings. The high and low values for the levels were determined, based on the prior experience of the BTU Engineering staff, to give a large representation of the operating space for the furnace. A total flow of 150 sccm was recommended so that the desired range of pressures could be achieved.

An independent set of experiments was conducted using flow through the load injector only. These experiments were conducted to investigate the parameters in the reaction model under conditions of high concentration variation. These experiments are discussed in Section 5.3.

Procedure

The experiments were conducted in a commercial BTU Engineering/Bruce Systems 7351C horizontal, hot-wall reactor with three zone temperature control. The equipment consisted of a 230/240 mm x 88.5 inch quartz process tube fitted with a 214/220 mm x 79 inch quartz liner. The furnace had a three zone heater with a 32 inch flat zone. The total wafer load consisted of 150 six inch wafers, positioned in six quartz boats, each holding 50 wafers and being 6 inches long with 3/32 inches center to center spacing. Only 25 wafers were inserted per boat giving a wafer spacing of 3/16 inches

center to center. The first and last boats were dummy boats. The production load thus consisted of 100 wafers in the center four boats.

The recipe used to run the furnace was the recipe BTU recommends for furnace operation when running a flat polysilicon process. The experiments were run by varying parameters in the recipe according to the L_9 experiment design structure. The deposition time in each experiment was 75 min. Thirteen test wafers were inserted at locations 20, 26, 35, 45, 55 65, 75, 85, 95, 105, 115, 124, and 130 in the wafer load of 150 wafers (wafers 20 and 130 were in the dummy load).

These wafters were then measured with a Nanometrics/AFT 010-0180 nanospec and a Gaertner ellipsometer. Measurements were made at the top, bottom, center, left, and right of the wafers. The average of these measurements was the growth rate attributed to each of the wafers.

Experimental Process Optimization

The experiment factors from the L_9 design were optimized according to Taguchi's signal-to-noise ratio (SN) criteria for "nominal is best".

$$SN = 10 \log \left(\frac{\text{mean}^2}{\text{variance}} \right) \tag{15}$$

The SN is calculated by determining the mean of the axial deposition profile and calculating the variance of the measured data from this mean. The "nominal is best" SN is used because the optimum deposition profile would be flat with the least variance about the mean. This criteria for the SN gives a measure of the relative flatness of the profile. The assumption is that the mean can be scaled to the desired value. In this case, the mean can be scaled by time to obtain the desired amount of total deposition.

Table 1 shows the calculated signal-to-noise ratios for the L_9. Figure 7 and Figure 8 show the profiles from experiments 1 and 9 of the experimental array, indicating the range of results attained.

Choice of the optimum factor level settings was determined by averaging the SN for the experiments in which the particular factor level was used. For example, the SN for experiments 1, 2, and 3 were averaged to obtain the relative SN attributed to factor 1, level 1. The optimum settings were the factor levels with the largest SN's. The optimum parameter levels were found to be the lowest pressure (200 mtorr), the middle land injector flow (60 sccm), the lowest center injector flow (40 sccm) and the middle

injector position 22 cm right of center). Figure 9 shows the corresponding deposition profile using these settings. This optimum profile is more axially uniform than the previous L_9 experiments and the base line, thus indicating that the equipment was optimized.

5. CALIBRATION OF CONSTANTS

Parameter Fit to L_9 array.

The kinetic constants in the reaction rate model, Equation 5, and the injector constant, Equation 7, were fit using a design optimization package called OPTDES developed at Brigham Young University [13]. This package used a non-linear reduced gradient method to optimize a least square objective function. The objective function included all nine experiments by summing the squared differences between the 13 measured data values and the predicted values at the corresponding wafer locations for each individual experiment in the L_9 and then summing the values for all nine experiments.

The constants obtained for the reaction model were:

$$k_1 = 1.202\text{x}10^{10} \exp(-18,500/T)\ \text{mol/m}^2\text{/s/atm}$$
$$K_e = 0.386\text{x}10^{5}\ \text{atm}^{-1}$$
$$K_h = 1.904\text{x}10^{4}\ \text{atm}^{-1/2}$$

The kinetic constants found from this regression are substantially different from those obtained by Roenigk and Jensen[6]:

$$k_1 = (1.6 \pm 0.4)\text{x}10^{9} \exp(-18,500/T)\ \text{mol/m}^2\text{/s/atm}$$
$$K_e = (0.7 \pm 0.1)\text{x}10^{5}\ \text{atm}^{-1}$$
$$K_h = (0.6 \pm 0.3)\text{x}10^{2}\ \text{atm}^{-1/2}$$

Our kinetic constants indicate that hydrogen has a much greater effect in inhibiting the reaction rate; K_h is three orders of magnitude larger. The Arrhenius constant in k_1 is larger by an order of magnitude due to the large increase of K_h. We also found that K_e has relatively little effect on the predicted profiles in the range of 0.01 x 10^5 atm^{-1} to 1.0 x 10^5 atm^{-1} since the denominator is dominated by the effect of hydrogen. The activation energy in k_1, Equation 6, was not fit to the data, but was determined by other investigators [14].

The ramp injector gave slightly better maximization of the objective function criteria. The injector fitting constant of Equation 7 was found to be 5.165. To give an idea of what this means, for an injector flow rate of 60 std cm^3/min the modeled injected gas length is approximately 3 cm.

Choice of Injector Model

The choice of the model to represent the gas spray dynamics from the two moveable injectors was done based on two criteria: a qualitative analysis of the injection dynamics; and maximization of the least square objective function.

After the model parameters were fit to the L_9 data it was found that the injector constant for the flat injector function was approximately twice that of the ramp function: *9.626 (flat) v. 5.165 (ramp)*. This is consistent with the idea that the flat model represented an average approximation of the ramp model. The ramp function gave a slightly larger objective function result, indicating that the variance from the measured data was less for this function.

Since qualitatively the ramp function is more representative of the expected gas dynamics and the objective function criteria was maximized for the ramp function, the ramp function was chosen as the injector model.

Calibration Check

To check the use of the L_9 experiment set for calibration of the kinetic parameters a "mini-experiment" matrix was designed to capture the main effects of the kinetic parameters in Equation 5. The design consisted of three experiments in which the pressure and temperature were varied. In these experiments, the total flow was introduced through the load injector only. This was done in order to maximize the effect of the silane depletion and hydrogen generation on the deposition profile. The constants found in the parameter fit to this data are:

$$k_1 = 1.412\text{x}10^{10} \exp(-18,500/T)\ \text{mol/m}^2\text{/s/atm}$$

$$K_s = 0.368\text{x}10^{6}\ \text{atm}^{-1}$$

$$K_h = 1.814\text{x}10^{4}\ \text{atm}^{-1/2}$$

These constants are very close to those obtained by the parameter fit to the L_9 design, thus, indicating that the model can be extrapolated quite effectively beyond its range of calibration.

6 MODEL RESULTS

The predicted deposition profiles closely resemble the measured profiles. Figures 10 - 18 give the predicted deposition profiles for each of the L_9 experiments and the optimized parameter run according to the previous Taguchi analysis with the measured data.

The primary purpose of our model is to predict the thickness variation down the length of the tube. Given a combination of process parameters (flow rates, injector positions, pressure, temperature, etc.) a calibration experiment can be run to determine the amount of time needed to determine a target thickness. The predicted profiles are therefore normalized to the means of the respective runs. Normalization was achieved by shifting the entire predicted curve by a multiplicative constant equal to the ratio of the measured profile mean to the predicted profile mean for each profile. In the worst case, Experiment 1, the mean of the predicted profile was 10% lower than the mean of the experimental data. The average error in the prediction of the mean deposition rate was 5%. No systematic explanation was found for the deviation in the mean growth rates.

The peaks in the predicted profiles are due to the approximation of the injector gas mechanics. It should be noted that the predicted profiles give an accurate representation of the measured profiles except at positions localized above the injector exit positions.

As a measure of the accuracy of the model to predict the variations in the equipment parameters, the signal-to-noise ratios, according to Equation 15, of the measured profiles were plotted against the SN of the predicted deposition profiles, Figure 19. A perfect correspondence would have resulted in a line with a slope of one and a y-intercept of zero. The results (slope = 1.05, y-intercept = -7.1) indicate that the model accurately predicts the effect of variations in the equipment parameters on the axial deposition profile. The slope shows that the measured data and predicted data correlate very well since they follow the same trend. The y-intercept is a measure of a constant offset of the predicted data from the measured data. Meaning, in this case, that the model predicts higher SN for each of the experiments than those found from the measured data.

7 PROCESS OPTIMIZATION

To determine the most robust operating setting for the modeled equipment, disturbance factors were add to the model. The disturbance factors were introduced by adding an outer array [8] to the model, Figure 21. The outer array consisted of an orthogonal array which allowed a maximum of 15 noise factors. each at two levels.

Thirteen equipment parameters were chosen for inclusion of disturbances: 3 thermocouple locations, 3 thermocouple temperatures, 2 silane injector locations, 3 injector flow rates, pressure, and wafer load position. The disturbances ranged from ±5% for pressure to ±2% on injector location as indicated by the Δ's in Figure 21.

The process optimization consisted of using OPTDES, the design optimization package used for the parameter fits previously discussed, to find the set of operating parameters which gave the least variance about the mean for the sum of the 16 runs in the outer array. Referring to Figure 20, the flow of the optimization can be understood. OPTDES selects a set of equipment parameters, 2 flow rates (the total flow was constrained to 150 sccm), pressure, and the position of 2 injectors (center and source). The disturbances are imposed on the process parameters and other parameters in the equipment model. These inputs are passed to the equipment model, which delivers a predicted output deposition profile. The resulting profile is stored and the model iterates through each of the 16 runs to obtain the profile with each set of disturbances. After the 16 iterations, the mean of the profiles is calculated and the signal-to-noise ratio of all the profiles from these means is found and returned to OPTDES. From this information, OPTDES finds another set of operating points and continues until the largest signal-to-noise ratio is found, indicating the optimum set of equipment parameters.

The result of this process optimization is shown in Figure 21. Plotted along with the process optimization result are the plots of the experimentally determined optimum settings as found in Section 4.3. The settings determined by both of these methods are very close. This shows that the model with the disturbances accurately simulates the actual equipment.

8 CONCLUSION

A model has been developed which predicts the wafer to wafer deposition rate of polysilicon down the length of a horizontal tube furnace. Inputs to the model include: silane flow rates from three injectors, two injector locations, locations of and temperatures of three thermocouples, operating pressure, the number of wafers, wafer diameter, the location of the wafer load, and other physical dimensions of the furnace such as tube length and liner diameter. The model construction consists of a one-dimensional finite difference numerical representation of the convective and diffusive fluxes of silane and hydrogen. Silane is injected and hydrogen is generated by reaction at the wafers and tube wall. The silane injection and mixing was modeled with a ramp function which is an approximate model incorporating a qualitative understanding of the injection phenomena.

Parameters in the reaction kinetics model and injector function were fit to a set of nine statistically designed experiments which varied four parameters, two injector flow rates, one injector position, and pressure, over three levels. These parameters were fit to all nine experiments in the design to give the best possible fit to all the data. The reaction kinetic parameters were independently fit to a second set of three experiments in which flow was admitted through only the load injector. The experiments were designed to explore a different region of operation, spanning a wide range of silane and hydrogen concentrations. The close correspondence between the reaction parameters fit to each set of experiments demonstrated that the model extrapolates well to regions beyond the scope of the initial experimental space.

An extension of the model, which included equipment disturbances, was used to optimize the process by finding the settings at which the process was most robust; that is, the settings which gave the flattest profile over the full range of disturbance. The settings predicted by the model correspond very closely to an experimentally determined robust operating point as determined by a Taguchi signal-to-noise ratio analysis of the statistically designed experimental space.

For the process engineer, the model is well suited for process optimization. Equipment and process disturbances can also be included to determine the most robust operating point, as has been demonstrated. Furthermore, the model can be extended for use in on-line quality control. Having optimized the operating point by process optimization, the operator can use the model to correct the equipment settings, based on product measurements, to maintain the required deposition specifications.

The equipment designer will find the model useful for testing new ideas in equipment design. Minor extensions can be made to the model, such as adding injectors, to determine the benefits of these changes without the cost of materials and experimentation.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of the Defense Advanced Research Projects Agency under contract N0014-85-K-0213 as well as the support of the Microelectronics and Computer Technology Corporation.

The authors would like to thank BTU Engineering Corporation for providing advice, guidance, and access to experimental facilities.

References

[1] M.L. Hitchman, J. Kane, and A.E. Widmer, "Polysilicon Growth Kinetics in a Low Pressure Chemical Vapour Deposition Reactor", *Thin Solid Films*, vol. 59, pp. 231-247, 1979.

[2] C.H.J.V.D. Brekel and L.J.M. Bollen, "Low Pressure Deposition of Polysilicon Silicon From Silane," *Journal of Crystal Growth*, vol. 54, pp. 310-322, 1981.

[3] W.A.P. Claassen *et al.*, "The Deposition of Silicon from Silane in a Low-Pressure Hot-Wall System," *Journal of Crystal Growth*, vol. 54, pp. 259-266, 1982.

[4] S. Middleman and A. Yeckel, "A Model of the Effects of Diffusion and Convection on the Rate and Uniformity of Deposition in a CVD Reactor," *Journal of the Electrochemical Society*, vol. 133, no. 9, pp. 1051-1056, 1986.

[5] K.F. Jensen and D.B. Graves, "Model and Analysis of Low Pressure CVD Reactors," *Journal of Electrochemical Society*, vol. 130, no. 9, pp. 1950-1957, 1983.

[6] K.F. Roenigk and K.F. Jensen, "Analysis of Multicomponent LPCVD Processes," *Journal of the Electrochemical Society*, vol. 132, no. 2, pp. 448-454, 1985.

[7] K. Lin, *An Expert System For Polysilicon Recipe Generation*. Master's thesis, University of California, Berkeley, July 1987.

[8] G. Taguchi, *Introduction to Quality Engineering*. Tokyo: Asian Productivity Organization, 1986.

[9] G.E.P. Box, W.G. Hunter, and J.S. Hunter, *Statistics for Experimentors*. John Wiley & Sons, 1978.

[10] G. Birkhoff and E.H. Zarantonello, *Jets, Wakes and Cavities*. New York: Academic Press, Inc., 1957.

[11] G.P. Brown, A. DiNardo, et al., "The Flow of Gases in Pipes at Low Pressures," *Journal of Applied Physics*, vol. 17, pp. 802-813, October 1946.

[12] M.S. Phadke, R.N. Kackar, D.V. Speeney, and M.J. Grieco, "Off-Line Quality Control in Integrated Circuit Fabrication Using Experimental Design," The Bell System Technical Journal, vol. 62, no. 5, pp. 1273-1309, 1983.

[13] A Parkinson, R. Balling, and J. Free, "OPTDES: A Software System for Optimal Engineering Design," in *Proceedings of the American Society of Mechanical Engineers Conference on Computers and Engineering*, (Las Vegas, Nevada), August 1986.

[14] W.A. Bryant, "The Kinetics of the Deposition of Silicon by Silane Pyrolysis at Low Temperatures and Atmospheric Pressure," *Thin Solid Films*, vol. 60, pp. 19-25, 1979.

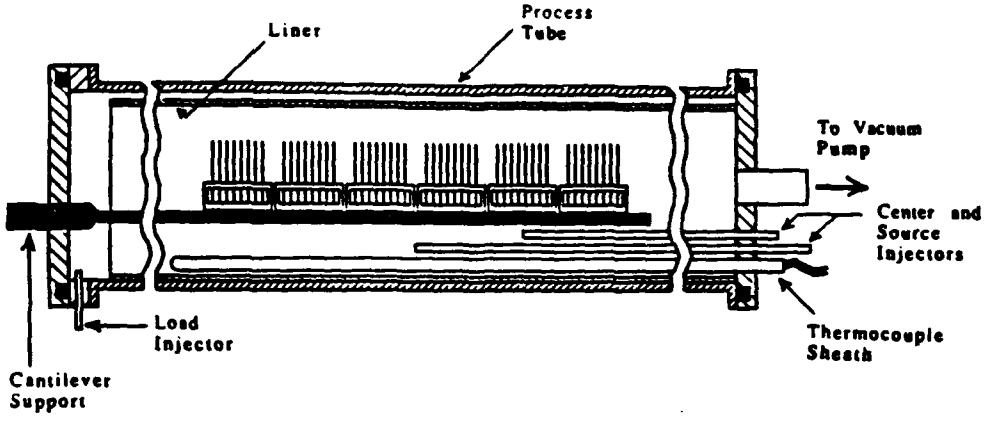

Figure 1. Schematic of LPCVD reactor.

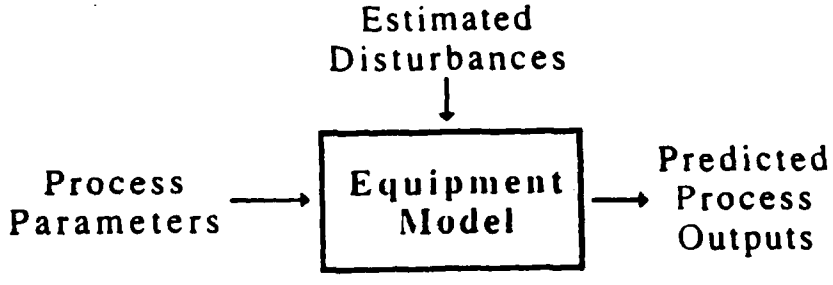

Figure 2. Generic representation of an equipment model.

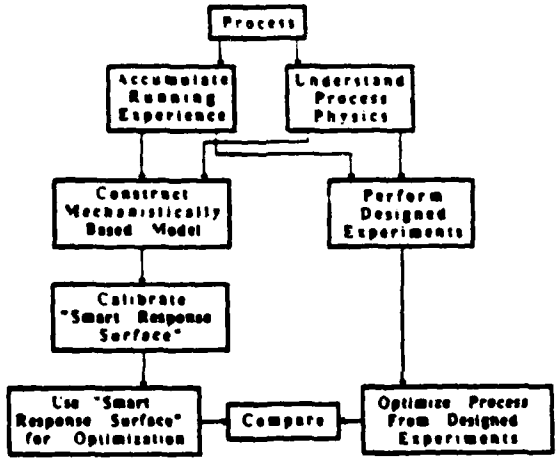

Figure 3. Flow chart of this paper.

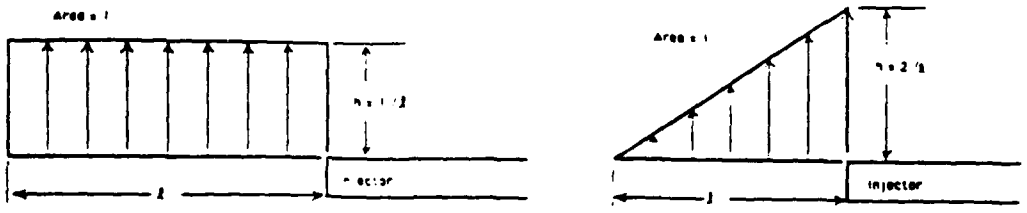

Figure 4. Representation of the source and center injector empirical formulations: a. Flat model. b. Ramp model.

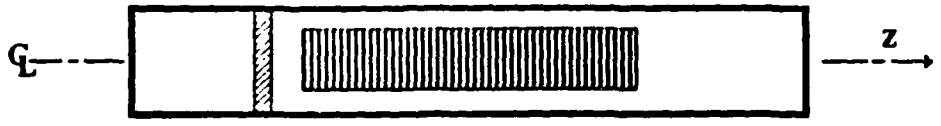

Figure 5. Schematic one-dimensional representation of the numerical model.

Experiment Number	Pressure (mtorr)	Q_{load} (% of total)	Q_{center} (% of total)	X_{source} (% of tube length from center)
1	200	20	26.7	9
2	200	30	36.7	12
3	200	40	46.7	15
4	250	20	36.7	15
5	250	30	46.7	9
6	250	40	26.7	12
7	350	20	46.7	12
8	350	30	26.7	15
9	350	40	36.7	9

$Q_{load} + Q_{center} + Q_{source} = Q_{total} = 150$ sind cm^3 / min

Process Temperature = 625 °C

Figure 6. Taguchi L_9 array.

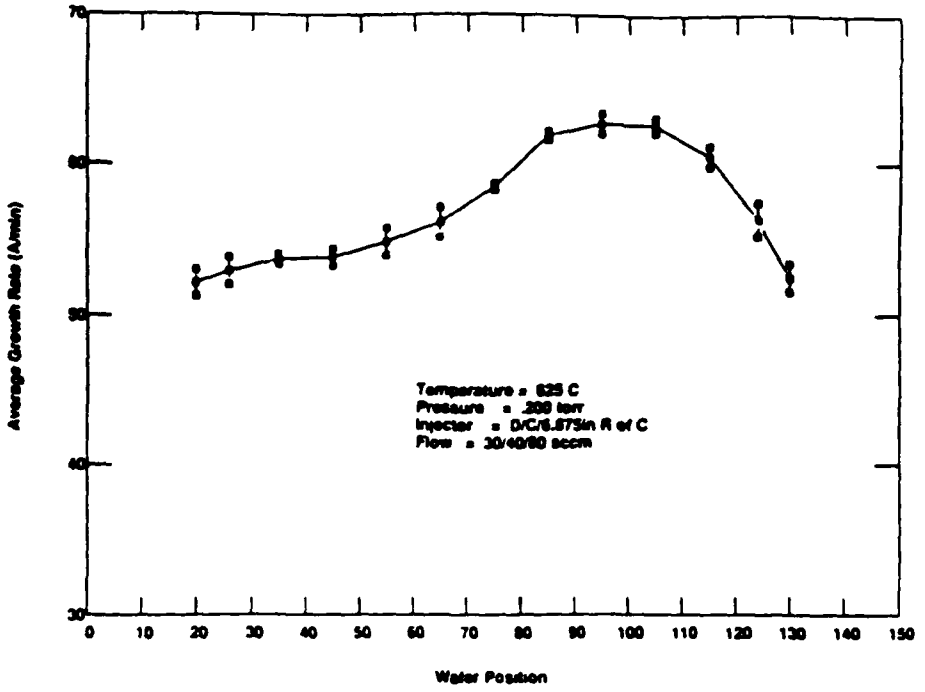

Figure 7. Deposition profile for experiment 1 of L_9 array.

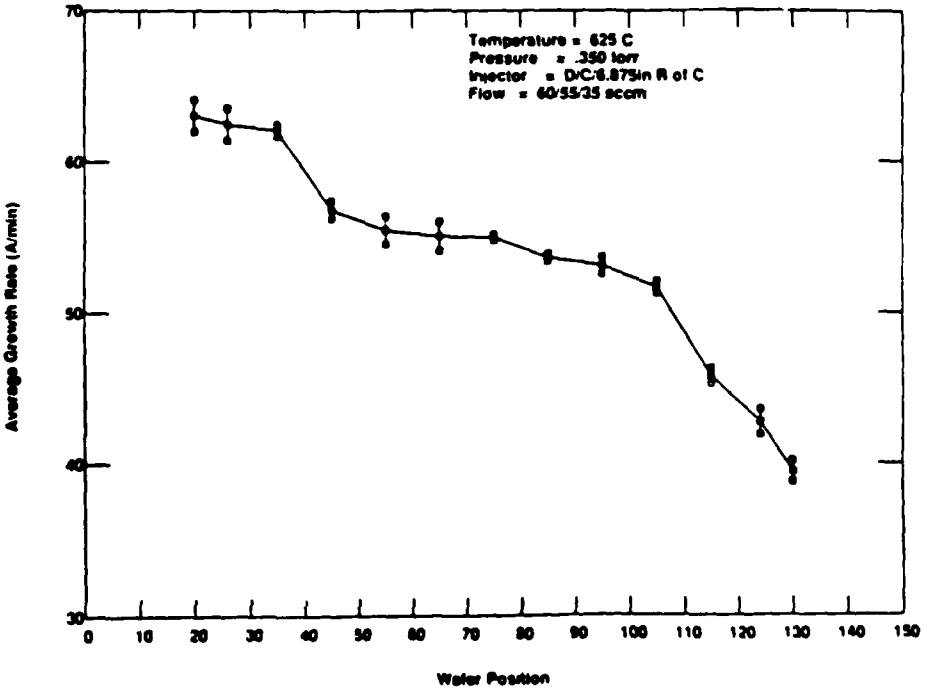

Figure 8. Deposition profile for experiment 9 of the L_9 array.

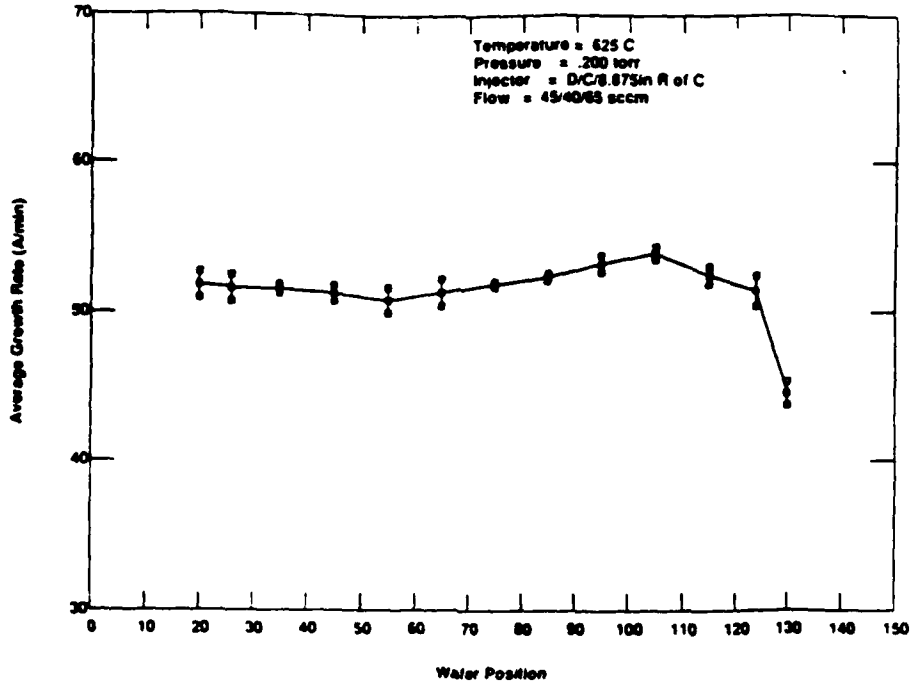

Figure 9. Deposition profile for experimentally determined optimum factor levels.

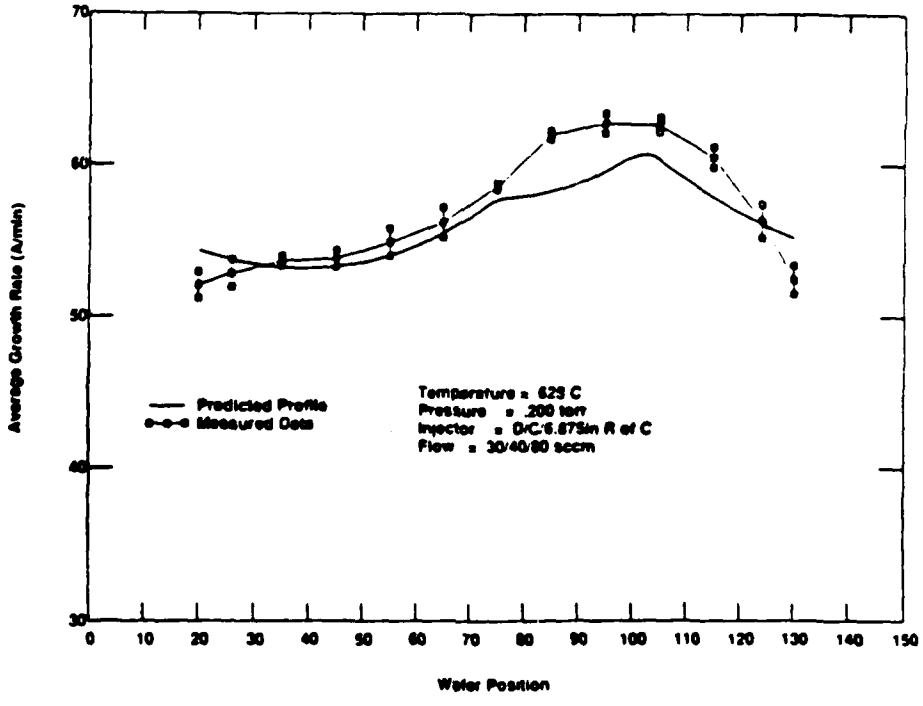

Figure 10. Growth rate profile and model prediction for experiment 1.

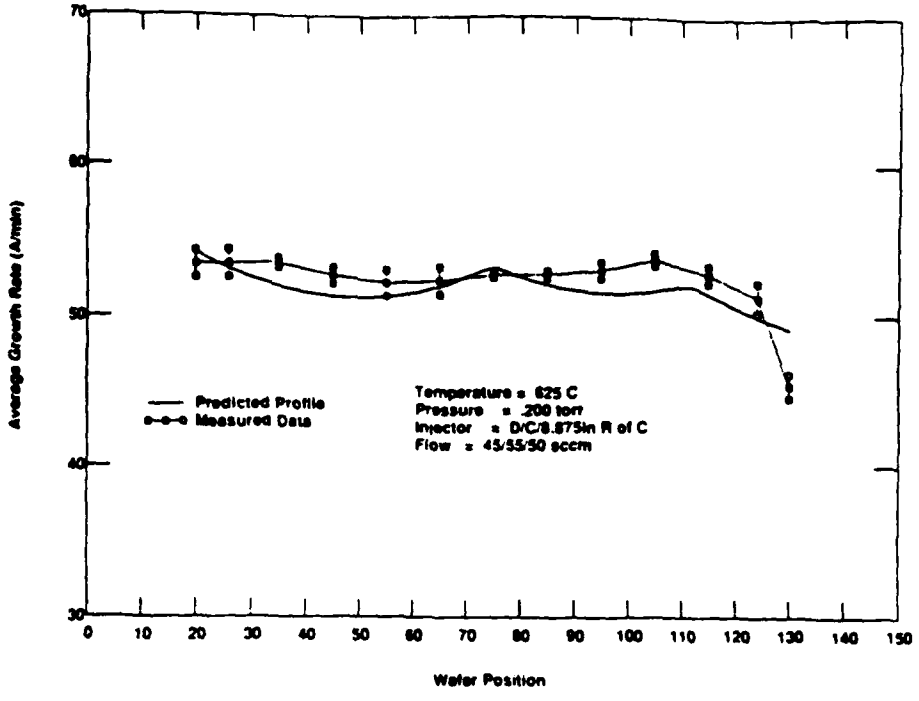

Figure 11. Growth rate profile and model prediction for experiment 2.

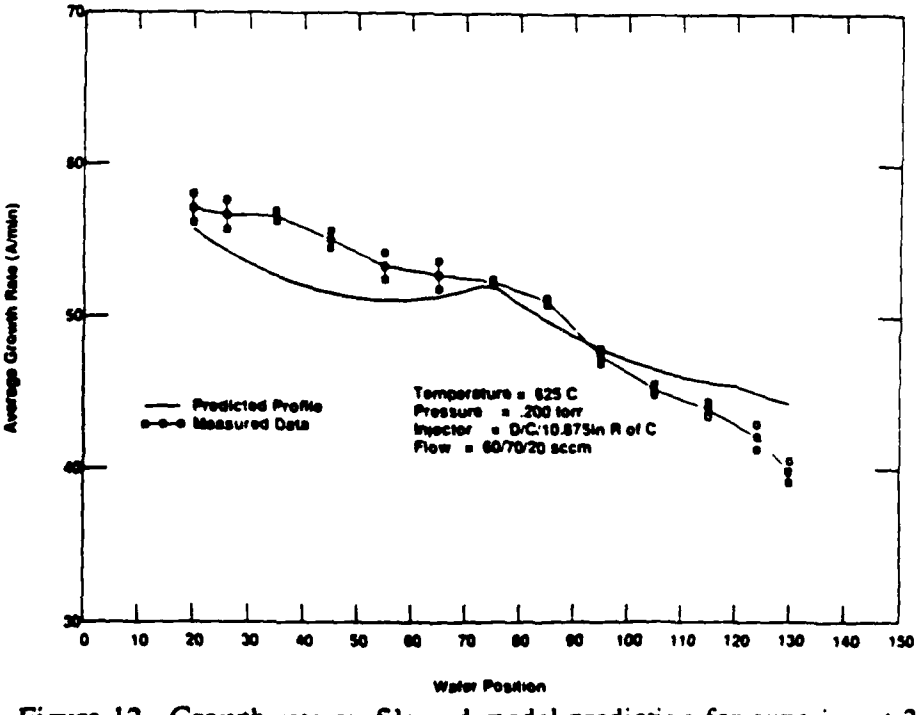

Figure 12. Growth rate profile and model prediction for experiment 3.

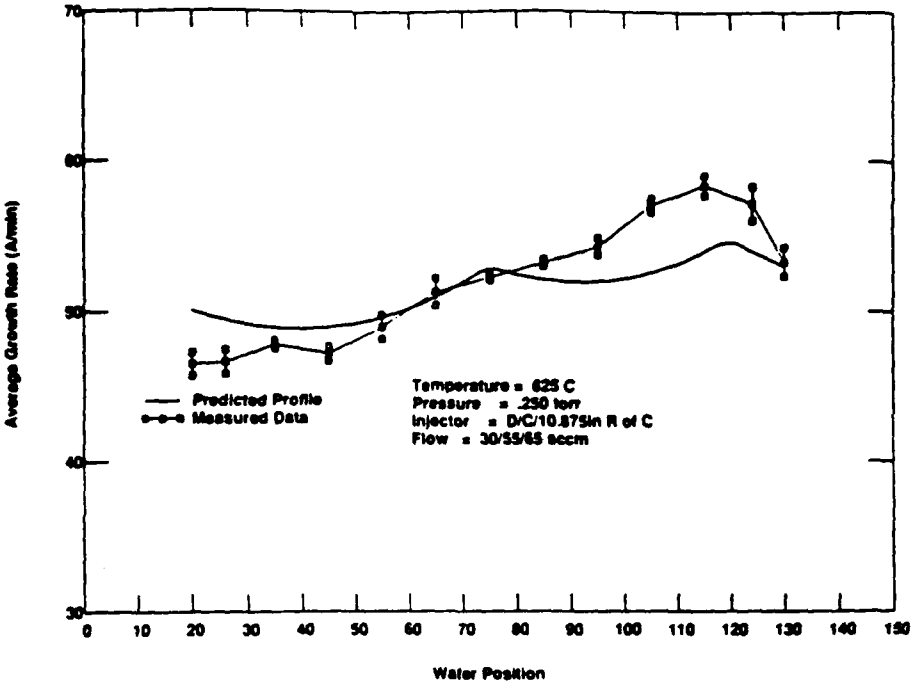

Figure 13. Growth rate profile and model prediction for experiment 4.

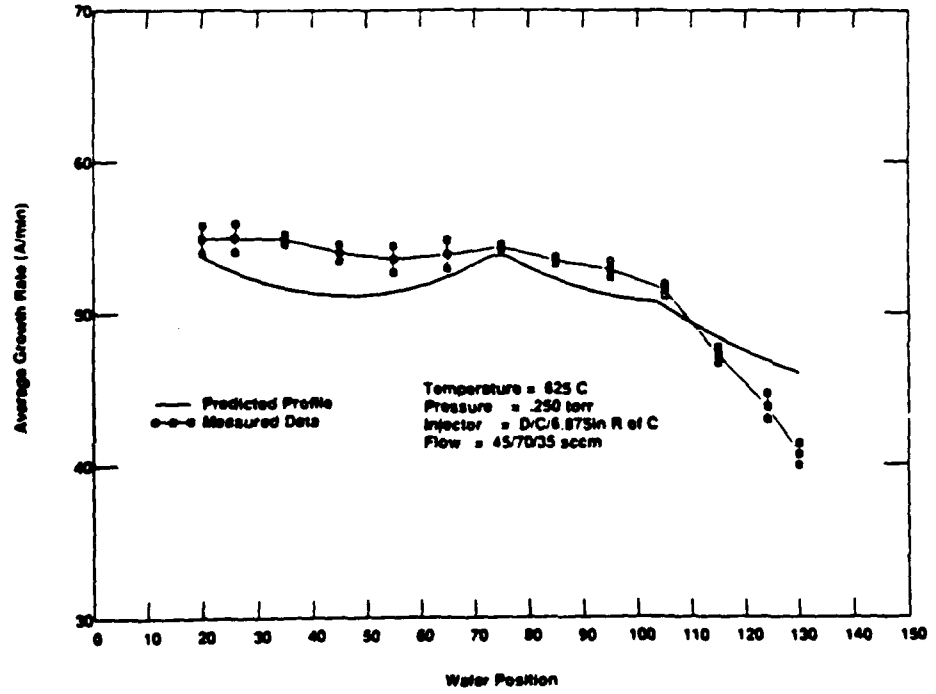

Figure 14. Growth rate profile and model prediction for experiment 5.

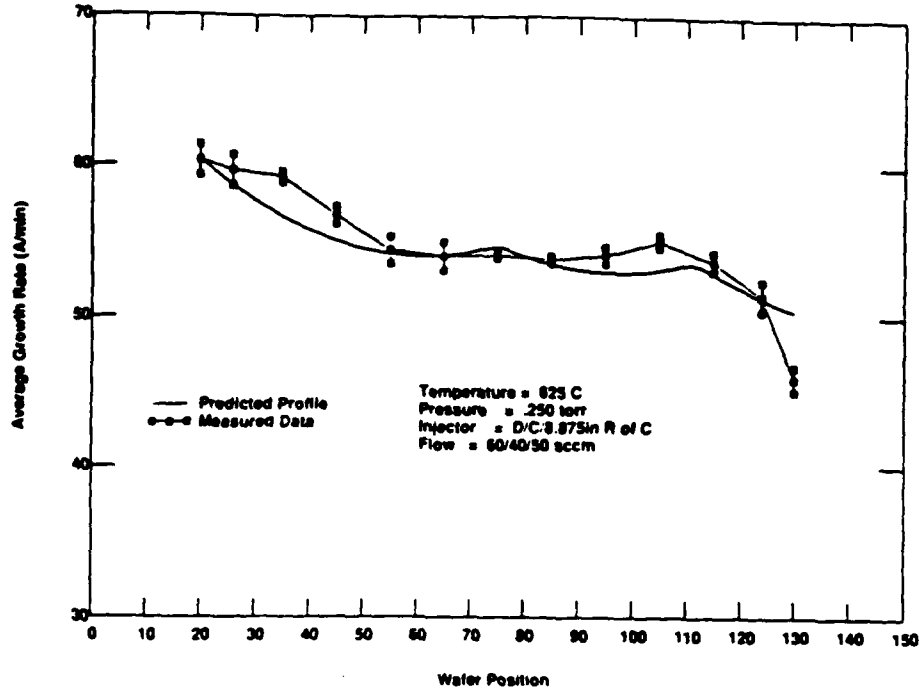

Figure 15. Growth rate profile and model prediction for experiment 6.

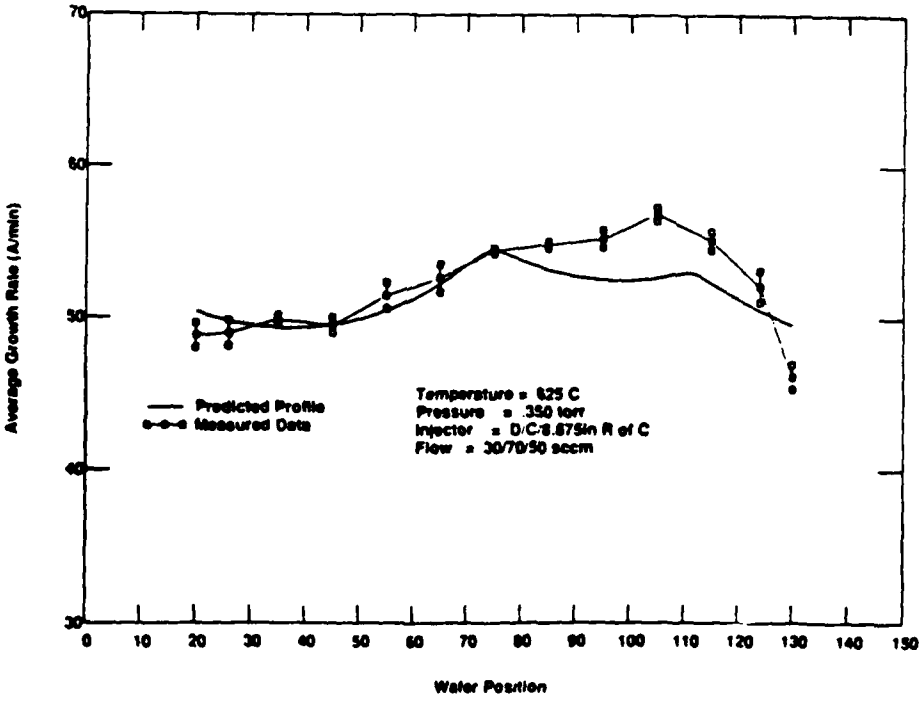

Figure 16. Growth rate profile and model prediction for experiment 7.

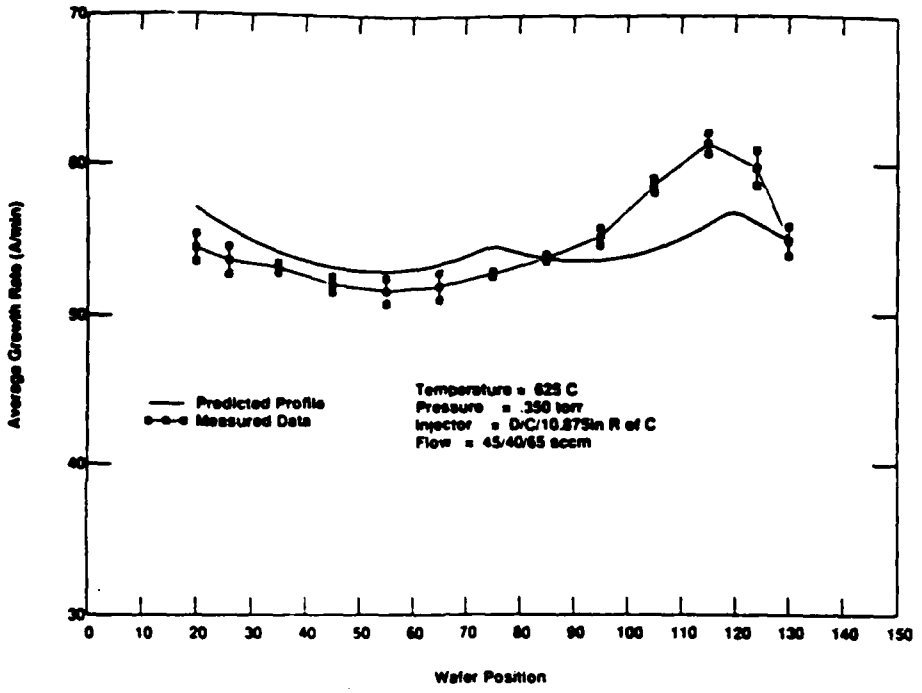

Figure 17. Growth rate profile and model prediction for experiment 8.

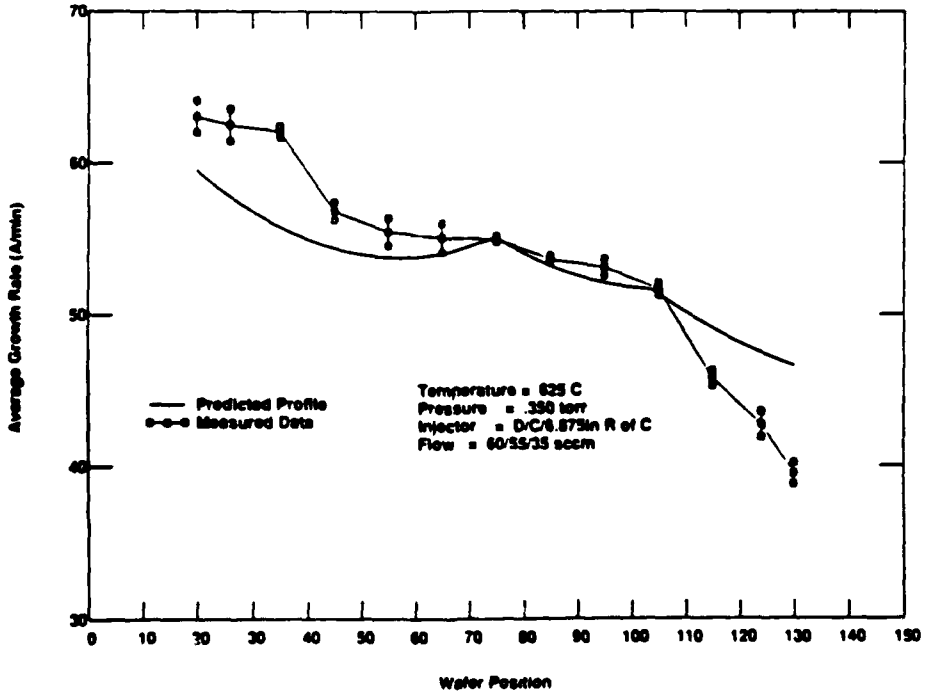

Figure 18. Growth rate profile and model prediction for experiment 9.

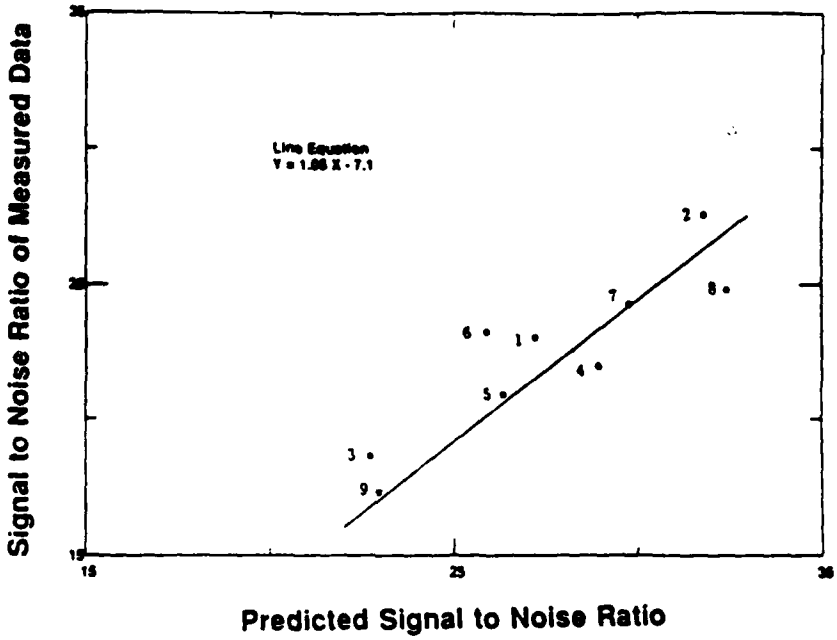

Figure 19. SN ratio predicted by model versus observed.

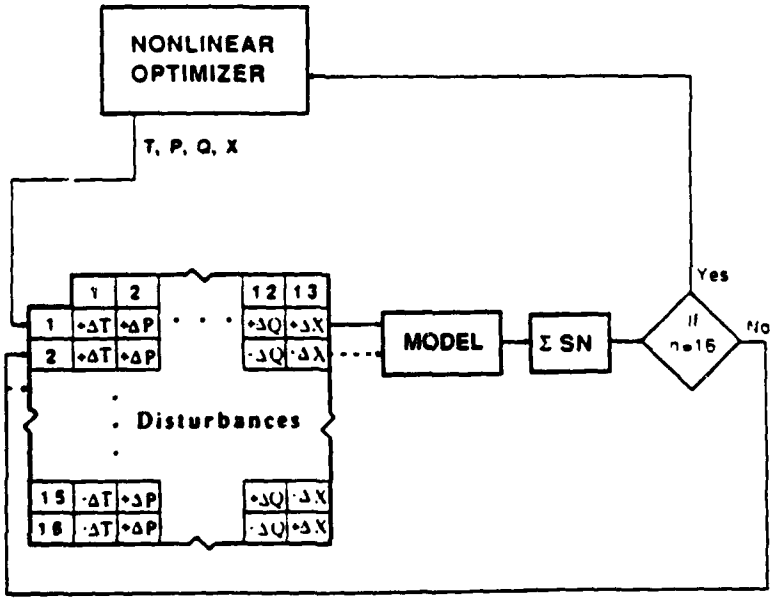

Figure 20. Optimization of profile using the model.

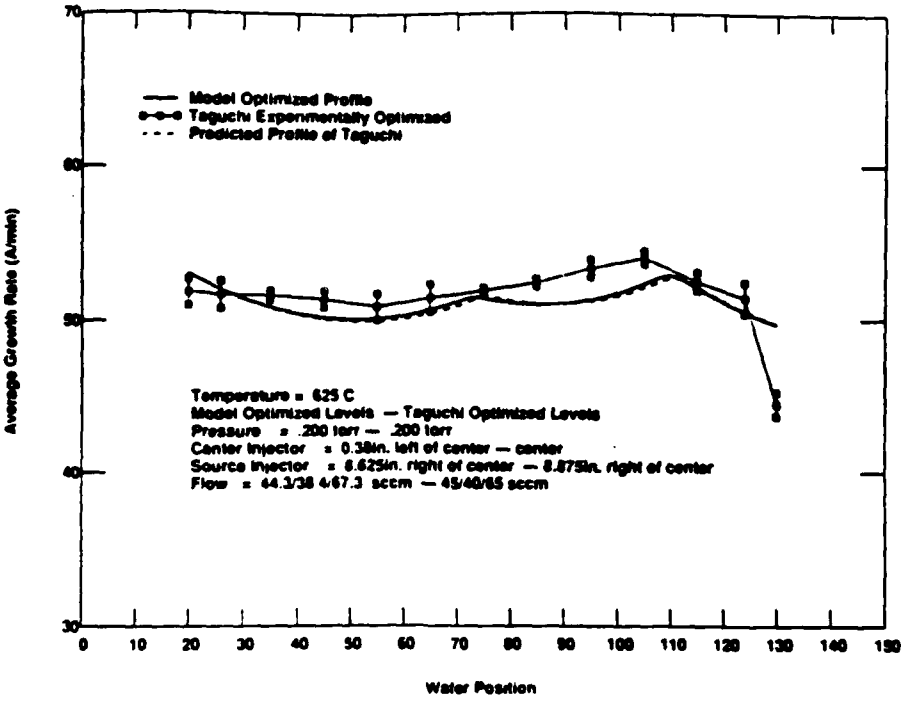

Figure 21. Result of optimization from model compared with optimization from experiments.

Runs	Parameters				S/N
	1	2	3	4	
1	1	1	1	1	23.01
2	1	2	2	2	27.55
3	1	3	3	3	18.61
4	2	1	2	3	21.91
5	2	2	3	1	20.89
6	2	3	1	2	23.24
7	3	1	3	2	24.23
8	3	2	1	3	24.79
9	3	3	2	1	17.30

Table 1. Signal-to-noise ratio results for L_9 design.